(54) Title: SYSTEMS AND METHODS FOR BIOPOLYMER ENGINEERING
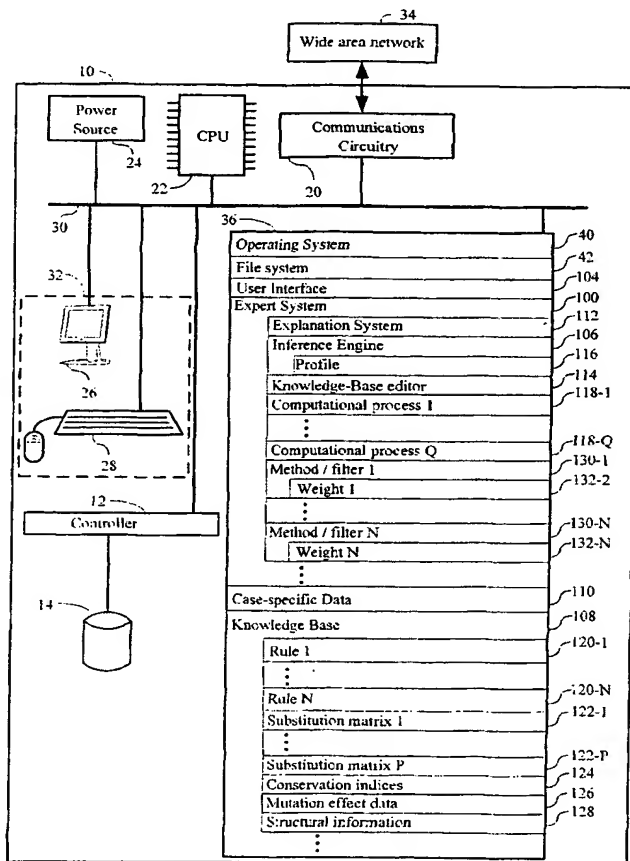
(57) Abstract: Methods, computer systems, and computer program products for biopolymer engineering. A variant set for a biopolymer of interest is constructed by identifying, using a plurality of rules, a plurality of positions in the biopolymer of interest and, for each respective position in the plurality of positions, substitutions for the respective position. The plurality of positions and the substitutions for each respective position in the plurality of positions collectively define a biopolymer sequence space. A variant set comprising a plurality of variants of the biopolymer of interest is selected. A property of all or a portion of the variants in the variant set is measured. A sequence-activity relationship is modeled between (i) one or more substitutions at one or more positions of the biopolymer of interest represented by the variant set and (ii) the property measured for all or the portion of the variants in the variant set. The variant set is redefined to comprise variants that include substitutions in the plurality of positions that are selected based on a function of the sequence-activity relationship.

GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI. CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.